The Annals of Probability
1989, Vol. 17, No. 1, 1–8

A SHARP DEVIATION INEQUALITY FOR THE STOCHASTIC TRAVELING SALESMAN PROBLEM

BY WANSOO T. RHEE[1] AND MICHEL TALAGRAND[2]

The Ohio State University and University of Paris VI and The Ohio State University

Let T_n denote the length of the shortest closed path connecting n random points uniformly distributed over the unit square. We prove that for some number K, we have, for all $t \geq 0$,

$$P(|T_n - E(T_n)| \geq t) \leq K \exp(-t^2/K).$$

1. Introduction. The famous traveling salesman problem (TSP) requires finding the length T_n of the shortest path connecting n points $X_1, \dots, X_n$ of the plane, that is, the infimum over all permutations σ of $\{1, \dots, n\}$ of the quantity

$$\|X_{\sigma(n)} - X_{\sigma(1)}\| + \sum_{1 \leq i < n} \|X_{\sigma(i+1)} - X_{\sigma(i)}\|.$$

We are here concerned with a stochastic version of the problem, where the points $X_1, \dots, X_n$ are independent and distributed uniformly on the unit square $[0,1]^2$. A number of recent papers [2–4] have proved that the random variable T_n is remarkably concentrated around its mean. The objective of the present paper is to prove the inequality

$$P(|T_n - E(T_n)| > t) \leq K \exp(-t^2/K). \tag{1}$$

In order to make our results applicable to other problems of geometric probability (e.g., the length of a Steiner tree or a rectilinear Steiner tree through $X_1, \dots, X_n$) and to isolate the properties of the TSP that we really need, it is suitable to state a more general result. Suppose that to each finite subset F of the unit square we associate a number $f(F)$, such that for each finite subset F and each $x \in [0,1]^2$

$$f(F) \leq f(F \cup \{x\}) \leq f(F) + d(x, F), \tag{2}$$

where $d(x, F) = \min\{d(x, y)\colon\ y \in F\}$. Consider independent random variables $X_1, \dots, X_n$ uniformly distributed over the unit square and let $U_n = f(\{X_1, \dots, X_n\})$.

THEOREM 1. *There exists a number K, independent of f and n, such that for all $t \geq 0$,*

$$P(|U_n - E(U_n)| > t) \leq K \exp(-t^2/K). \tag{3}$$

Received October 1987; revised March 1988.
[1]Research supported in part by NSF Grant DCR-86-01025.
[2]Research supported in part by NSF Grant DMS-86-03951.
AMS 1980 *subject classifications*. Primary 68C25; secondary 90C10, 65K05, 60G48.
Key words and phrases. Martingale inequalities, shortest path, exponential tail, stochastic model.

By scaling, we see that if (2) is replaced by

$$f(F) \le f(F \cup \{x\}) \le f(F) + ad(x, F) \tag{4}$$

for some constant $a > 0$, then

$$P(|U_n - E(U_n)| \le t) \le K \exp\left(-\frac{t^2}{a^2 K}\right). \tag{5}$$

In particular, the shortest tour through all the points of F satisfies (4) for $a = 2$.

It is easy to see that for t of the order $\sqrt{n}$, inequality (1) is optimal. Indeed, if H denotes the event $\{\forall\ i \le n,\ X_i \in [0, \frac{1}{2}]^2\}$, then $P(H) = 4^{-n}$. We have $E(T_n|H) = \frac{1}{2}E(T_n)$ by homogeneity. We know that for some constant c independent of n, we have $E(T_n) \ge 4c\sqrt{n}$. Thus

$$\begin{aligned} P\big(T_n \le E(T_n) - c\sqrt{n}\big) &\ge P\big(H \cap \{T_n \le E(T_n) - c\sqrt{n}\}\big) \\ &= P(H)P\big(T_n \le E(T_n) - c\sqrt{n}|H\big) \\ &= 4^{-n}P\big(T_n \le 2E(T_n) - 2c\sqrt{n}\big) \end{aligned}$$

by homogeneity and thus, by (1),

$$\begin{aligned} P\big(T_n \le E(T_n) - c\sqrt{n}\big) &\ge 4^{-n}P\big(\{T_n \le E(T_n) + 2c\sqrt{n}\}\big) \\ &\ge 4^{-n}\left(1 - K\exp\left(-\frac{4c^2 n}{K}\right)\right) \ge K' \exp\left(-K'\left(c\sqrt{n}\right)^2\right) \end{aligned}$$

for some constant K' independent of n.

We do not know whether inequality (3) is optimal when $t \ll \sqrt{n}$.

2. Plan of proof. As in our previous work on the topic, our method heavily relies on martingale difference sequence methods (m.d.s.). In [3] we obtained a bound

$$P(|T_n - E(T_n)| \ge t) \le 2\exp\big(-\alpha t^2/\log(1 + t)\big)$$

for some $\alpha > 0$ by interpolating between two known martingale inequalities. In the present work, we will use only standard martingale inequalities, but we will need a more detailed analysis and we will make use of a seemingly new principle, that may be of independent interest. It is explained after Lemma 1. Let $\mathscr{F}_i$ denote the σ-field generated by $X_1, \dots, X_i$. For simplicity, we denote by E^i the conditional expectation with respect to $\mathscr{F}_i$. Denoting by $[x]$ the integer part of x, we set $m = [n/2]$. In order to simplify notation, we will denote by K a universal constant that may vary from line to line.

LEMMA 1. *We have, for all* $t \ge 0$,

$$P(|E^m(U_n) - E(U_n)| \ge t) \le 2\exp(-t^2/K).$$

PROOF. The proof is almost identical to that of [2], Proposition 6, but we give it for completeness. We consider the m.d.s. associated to U_n, that is given by

$d_i = E^i(U_n) - E^{i-1}(U_n)$, so that $E^m(U_n) - E(U_n) = \sum_{1 \le i \le m} d_i$. We note, as in [2], Corollary 5, that $\|d_i\|_\infty \le K(n-i+1)^{-1/2}$, so that $\|d_i\|_\infty \le Kn^{-1/2}$ for $i \le m$. Finally, we conclude by using Azuma's inequality ([5], Lemma 4-2-3 and Exercise 4-2-2)

$$P\left(\left|\sum_{i \le m} d_i\right| \ge t\right) \le 2\exp\left(-t^2 \Big/ \left(2\sum_{i \le m} \|d_i\|_\infty^2\right)\right). \qquad \square$$

If one tries to apply the above method to obtain a bound for $P(|U_n - E(U_n)| > t)$, one obtains only a bound $2\exp(-t^2/K \log n)$, as in [2]. This failure is however due to the terms d_i for i close to n and this motivates Lemma 1. The idea is that we are left to control $P(|U_n - E^m(U_n)| > t)$. This can be done conditionally on $X_1, \dots, X_m$. But knowing $X_1, \dots, X_m$ provides precious information (at least for most of the choices of $X_1, \dots, X_m$) on $f(X_1, \dots, X_m, X_{m+1}, \dots, X_n)$. The success of that program is described in the following lemma, that is the essential part of the proof of Theorem 1.

LEMMA 2. *For each $t > 0$, there exists a subset H_t of $[0,1]^{2m}$ with the following properties*:

1. $P((X_1, \dots, X_m) \in H_t) \le K\exp(-t^2/K)$.
2. *If we define the random variable h by $h = f(\{x_1, \dots, x_m, X_{m+1}, \dots, X_n\})$, for $(x_1, \dots, x_m) \notin H_t$, we have*

$$P(|h - E(h)| \ge t) \le 2\exp(-t^2/K).$$

We conclude now the proof of Theorem 1. We observe that for $(x_1, \dots, x_m) \notin H_t$, the relation

$$P(|h - E(h)| > t) \le 2\exp(-t^2/K)$$

means that

$$P(|U_n - E^m(U_n)| > t | X_1 = x_1, \dots, X_m = x_m) \le 2\exp(-t^2/K).$$

It follows that

$$\begin{aligned} P(|U_n - E^m(U_n)| > t) &\le 2\exp(-t^2/K) + P((X_1, \dots, X_m) \in H_t) \\ &\le K\exp(-t^2/K). \end{aligned}$$

The result follows since

$$\begin{aligned} &P(|U_n - E(U_n)| > t) \\ &\quad \le P(|U_n - E^m(U_n)| > t/2) + P(|E^m(U_n) - E(U_n)| > t/2). \end{aligned}$$

3. Proof of Lemma 2. We first observe a simple property of f. Given two finite sets F and G in the unit square, let us call a F-spanning graph any graph that consists of line segments between points of $F \cup G$ and that contains a path from any point of G to at least a point of F. From condition (2), we deduce the following by induction over card G.

LEMMA 3. *If there is an F-spanning graph of length L in $F \cup G$, we have*

$$f(F) \le f(F \cup G) \le f(F) + L.$$

It is well known that a subset F of the unit square of cardinality n has a closed tour of length $\le 2\sqrt{n}$. So, if $x \in F$, we have

$$f(\{x\}) \le f(F) \le f(\{x\}) + 2\sqrt{n}. \tag{6}$$

On the other hand, for $x,\ y \in [0,1]^2$, we have

$$|f(\{x\}) - f(\{y\})| \le |f(\{x, y\}) - f(\{x\})| + |f(\{x, y\}) - f(\{y\})| \le 2\sqrt{2}.$$

It then follows from (6) that for any set F with card $F \le n$, we have $a \le f(F) \le a + 2(\sqrt{n} + \sqrt{2})$ for $a = \inf\{f(\{x\});\ x \in [0,1]^2\}$. Thus we have $a \le f(F) \le b$, where a, b are independent of F and $b - a \le K\sqrt{n}$. This shows that it is enough to prove Lemma 2 when $t \le \beta\sqrt{n}$, where β is some fixed number. Obviously, we can also assume $t \ge 1$.

For $k \ge 1$, we denote by $\mathscr{A}_k$ the natural collection of the 2^{2k} closed squares of side 2^{-k} that cover $[0,1]^2$. We denote by p the largest integer for which $2^{-p} \ge 1/t$. The computational part of the proof of Lemma 2 is contained in the following lemma, that will be proved in the next section.

LEMMA 4. *For $1 \le t \le \sqrt{n}$, there exists a subset H_t of $[0,1]^{2m}$ with the following properties*:

1. $P((X_1, \dots, X_m) \in H_t) \le K \exp(-t^2/K)$.
2. *For $(x_1, \dots, x_m) \notin H_t$, consider the union Z of all the squares of $\mathscr{A}_p$ that contain at least a point x_i, $i \le m$. Then $P(X_1 \notin Z) \le Kt^2/n$ and if we set $\phi(x) = d(x, \{x_1, \dots, x_m\})1_Z(x)$, we have $E(\phi^2(X_1)) \le K/n$.*

The idea is that condition $E(\phi^2(X_1)) \le K/n$ will provide the crucial control over the points in Z, as will be shown in Lemma 6. But we proceed first to show that the condition $P(X_1 \notin Z) \le Kt^2/n$ implies that the points that do not belong to Z are unimportant. We fix $(x_1, \dots, x_m) \notin H_t$. We set

$$v = f(\{x_1, \dots, x_m, X_{m+1}, \dots, X_n\} \cap Z). \tag{7}$$

The next lemma shows that v is a small perturbation of h and thus that it will be sufficient to study v instead of h.

LEMMA 5. *For some universal constant K_0, we have*

$$v \le h \le v + K_0 t.$$

PROOF. For $k \le p$, consider the collection J_k of squares of $\mathscr{A}_k$ that contain no points (x_i), $i \le m$. Denote by I_k the collection of squares of J_k that are contained in no square of J_l for any $l < k$. Let

$$F = \{x_1, \dots, x_m, X_{m+1}, \dots, X_n\} \cap Z$$

and

$$G = \{x_1, \dots, x_m, X_{m+1}, \dots, X_n\} \setminus Z.$$

Each point of $A \in I_k$ is within distance $\sqrt{2}\, 2^{-k+1}$ of a point of F. There is a tour through the points of $A \cap G$ of length $\leq 2^{-k+1}(\operatorname{card}(A \cap G))^{1/2}$. It follows that there is an F-spanning graph of $F \cup G$ of length L such that

$$L \leq \sum K 2^{-k}\left(1 + (\operatorname{card}(A \cap G))^{1/2}\right),$$

where the summation is over all $k \leq p$ and $A \in I_k$. Set $Y_k = \bigcup\{A;\ A \in J_k\}$. Using the Cauchy–Schwarz inequality, we have

$$\sum_{A \in I_k} (\operatorname{card}(A \cap G))^{1/2} \leq (\operatorname{card} I_k)^{1/2}(\operatorname{card}(Y_k \cap G))^{1/2}$$

so that, by Cauchy–Schwarz again

$$L \leq \sum_{k \leq p} K 2^{-k} \operatorname{card} I_k + K 2^{-k}(\operatorname{card} I_k)^{1/2}(\operatorname{card}(Y_k \cap G))^{1/2}$$

$$\leq K 2^p\left(\sum_{k \leq p} 2^{-2k} \operatorname{card} I_k\right) + K\left(\sum_{k \leq p} 2^{-2k} \operatorname{card} I_k\right)^{1/2} (\operatorname{card} G)^{1/2}.$$

Since $\sum_{k \leq p} 2^{-2k} \operatorname{card} I_k = P(X_1 \notin Z) \leq K t^2/n$, we have

$$L \leq K 2^p t^2/n + K(t^2/n)^{1/2} n^{1/2} \leq K(2^p + t) \leq Kt$$

by definition of p. The result then follows from Lemma 3, since $v = f(F) \leq f(F \cup G) = h \leq f(F) + L = v + L$. □

It follows from Lemma 5 that

$$P(|h - E(h)| \geq (2K_0 + 1)t) \leq P(|v - E(v)| \geq t).$$

Hence [changing t into $(2K_0 + 1)t$] Lemma 2 is a consequence of the following.

LEMMA 6. *If v is given by* (7), *we have*

$$P(|v - E(v)| \geq t) \leq 2\exp(-t^2/K).$$

PROOF. It relies once more on m.d.s. Let $d_i = E^i(v) - E^{i-1}(v)$, so that $v - E(v) = \sum_{m < i \leq n} d_i$. For $i > m$, let

$$v_i = f(\{x_1, \dots, x_m, X_{m+1}, \dots, X_{i-1}, X_{i+1}, \dots, X_n\} \cap Z).$$

If $X_i \notin Z$, we have $v_i = f$. Otherwise, from (2), we have

$$v_i \leq v \leq v_i + d(X_i, \{x_1, \dots, x_m\})$$

so that we have $v_i \le v \le v_i + \phi(X_i)$, and hence

$$E^i(v_i) \le E^i(v) \le E^i(v_i) + \phi(X_i),$$
$$E^{i-1}(v_i) \le E^{i-1}(v) \le E^{i-1}(v_i) + E\big(\phi(X_i)\big).$$

Since v_i is independent of X_i, we have $E^i(v_i) = E^{i-1}(v_i)$, so that we have

$$|d_i| \le \phi(X_i) + E\big(\phi(X_i)\big).$$

If $x \notin Z$, we have $\phi(x) = 0$. If $x \in Z$, then x is within distance $\sqrt{2}\,2^{-p}$ of $\{x_1, \dots, x_m\}$. Thus $|\phi(x)| \le \sqrt{2}\,2^{-p} \le K/t$ and hence $\|\phi\|_\infty \le K/t$. It follows that $\|d_i\|_\infty \le K/t$. Also, we have $E(\phi^2(X_i)) \le K/n$ and thus $\|E^{i-1}(d_i^2)\|_\infty \le K/n$. (It is to obtain that crucial property that we reduced the study of h to that of v.) It is shown in [1], Proposition 3.1, that for a m.d.s. $(d_i)_{i \le n}$ such that $\|d_i\|_\infty \le M$, we have

$$P\left(\left\{\left|\sum_{i=1}^{n} d_i\right| \ge t\right\}\right) \le 2\exp\left\{-\frac{t}{2M}\operatorname{arcsinh}\left(\frac{Mt}{2\sum_{i=1}^{n}\|E^{i-1}\big(d_i^2\big)\|_\infty}\right)\right\}.$$

(This is a martingale version of Prokhorov's inequality.) This implies the result. □

4. Proof of Lemma 4. Since we try to obtain a smallness condition on ϕ, the obvious idea is to try to make $x_1, \dots, x_m$ rather uniformly spread. We define q as the largest integer for which $2^{-2q} \ge (1/m)\log(em/t^2)$. We fix a number α such that $x/2 \ge \log ex^2$ for $x \ge \alpha$ and we define r as the largest integer for which $m2^{-2r} \ge \alpha$. As we observed, it is enough to prove Lemma 3 when $t \le \beta\sqrt{n}$, where β is universal, so we can assume that $\alpha/m \le (1/m)\log(em/t^2) \le 1/t^2$, so that $p \le q \le r$.

We set

$$a_k = \big(t^2 + k - p + 1\big)2^{2k}/m, \quad \text{if } p \le k < q,$$
$$a_k = 2^{6k}/m^2, \quad \text{if } q \le k \le r.$$

For $p \le k \le r$ we set $s_k = [a_k]$ and

$$V_k = \big\{(x_1, \dots, x_m);\ \text{at least } s_k + 1 \text{ squares of } \mathscr{A}_k \text{ do not meet } \{x_1, \dots, x_m\}\big\}.$$

We set $H_t = \bigcup_{p \le k \le r} V_k$ and we fix $(x_1, \dots, x_m) \notin H_t$. Then

$$P(X_1 \notin Z) \le 2^{-2p}s_p \le t^2/m \le Kt^2/n.$$

We observe that if $\phi(X_1) > 2^{-k+1}$, then clearly X_1 belongs to a square of $\mathscr{A}_k$ that does not meet $\{x_1, \dots, x_m\}$. So

$$P\big(\phi(X_1) > 2^{-k+1}\big) \le 2^{-2k}s_k.$$

Since, as we already observed, $|\phi| \le 2^{-p+1}$, we have

$$E\big(\phi^2(X_1)\big) \le 2^{-2r+2} + \sum_{p<k\le r} 2^{-2k+4} P\big(\phi(X_1) > 2^{-k+1}\big)$$
$$\le 2^{-2r+4} + 6 \sum_{p<k<q} 2^{-2k}\big(t^2 + k - p + 1\big)/m + 4 \sum_{q\le k\le r} 2^{2k}/m^2$$
$$\le K\big(1/m + 2^{-2p}t^2/m + 2^{2r}/m^2\big) \le K/n$$

since, by definition of p and r we have $2^{-2r} \le K/m$, $2^{2r} \le Km$, $2^{-p}t \le 2$. So it remains only to prove that

$$P\big((X_1,\dots,X_m) \in H_t\big) \le K \exp\big(-t^2/K\big).$$

To simplify notation, fixing $p \le k \le r$, we set $N = 2^{2k}$, $s = s_k + 1$, so $s \ge a_k$. We have

$$P\big((X_1,\dots,X_m) \in V_k\big) \le \binom{N}{s}(1 - s/N)^m$$
$$\le (eN/s)^s e^{-sm/N} = \exp\big(-s\big(m/N - \log(eN/s)\big)\big).$$

CASE 1. $p \le k < q$. Then $eN/s \le e2^{2k}/a_k \le em/t^2$, so that

$$\log(eN/s) \le \log\big(em/t^2\big) \le m2^{-2q} \le 2^{-2k-2}m$$

since $k < q$. Since $m/N = 2^{-2k}m$, $s \ge a_k$, we have

$$P\big((X_1,\dots,X_m) \in V_k\big) \le \exp\big(-2^{-2k-1}ma_k\big) \le \exp\big(-\big(t^2 + k - p + 1\big)/2\big).$$

CASE 2. $q \le k \le r$. Then $eN/s \le e2^{2k}/a_k = e2^{-4k}m^2$. Since $k \le r$, we have $2^{-2k}m \ge \alpha$ by definition of r, so that $\log(e2^{-4k}m^2) \le m2^{-2k-1}$ by the choice of α and hence $\log(eN/s) \le m2^{-2k-1}$. It follows that

$$P\big((X_1,\dots,X_m) \in V_k\big) \le \exp\big(-2^{-2k-1}ma_k\big) = \exp\big(-2^{4k-1}/m\big).$$

To conclude the proof, we observe that

$$\sum_{k\ge p} \exp\big(-\big(t^2 + k - p + 1\big)/2\big) \le K \exp\big(-t^2/2\big)$$

and that

$$\sum_{k\ge q} \exp\big(-2^{4k-1}/m\big) \le K \exp\big(-2^{4q-1}/m\big).$$

Now we have

$$2^{4q}/m \ge m\big(\log em/t^2\big)^{-2} \ge t^2/K. \qquad \square$$

REFERENCES

[1] JOHNSON, W. B., SCHECHTMAN, G. and ZINN, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.* **13** 234–253.

[2] RHEE, W. T. and TALAGRAND, M. (1987). Martingale inequalities and NP-complete problems. *Math. Oper. Res.* **12** 177–181.

[3] RHEE, W. T. and TALAGRAND, M. (1988). Martingale inequalities, interpolation and NP-complete problems. *Math. Oper. Res.* To appear.

[4] STEELE, J. M. (1981). Complete convergence of short paths and Karp's algorithm for the TSP. *Math. Oper. Res.* **6** 374–378.

[5] STOUT, W. F. (1974). *Almost Sure Convergence*. Academic, New York.

FACULTY OF MANAGEMENT SCIENCE
THE OHIO STATE UNIVERSITY
1775 COLLEGE ROAD
COLUMBUS, OHIO 43210

UNIVERSITY OF PARIS VI
EQUIPE D'ANALYSE-TOUR 46
4 PLACE JUSSIEU
75230 PARIS CEDEX 05
FRANCE
AND
DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
231 WEST 18TH STREET
COLUMBUS, OHIO 43210